

UNITED STATES DISTRICT COURT
SOUTHERN DISTRICT OF NEW YORK

----- X
H. CRISTINA CHEN-OSTER, SHANNA :
ORLICH, ALLISON GAMBA and MARY :
DE LUIS, :
:

Plaintiffs, :

v. :

GOLDMAN SACHS & CO. and THE :
GOLDMAN SACHS GROUP, INC., :

Defendants. :
----- X

10 Civ. 6950 (AT) (RWL)

ORAL ARGUMENT REQUESTED

**OPPOSITION TO PLAINTIFFS' MOTION FOR SUMMARY JUDGMENT
ON *PRIMA FACIE* ELEMENT OF DISPARATE IMPACT CLAIMS**

Barbara B. Brown (*admitted pro hac vice*)
Carson H. Sullivan (*admitted pro hac vice*)
PAUL HASTINGS LLP
875 15th Street, N.W.
Washington, D.C. 20005

Patrick W. Shea
PAUL HASTINGS LLP
200 Park Avenue
New York, New York 10166

Robert J. Giuffra, Jr.
Sharon L. Nelles
Ann-Elizabeth Ostrager
Hilary M. Williams
SULLIVAN & CROMWELL LLP
125 Broad Street
New York, New York 10004

Amanda Flug Davidoff
SULLIVAN & CROMWELL LLP
1700 New York Avenue, N.W., Suite 700
Washington, D.C. 20006

*Attorneys for Defendants
Goldman Sachs & Co. and
The Goldman Sachs Group, Inc.*

October 12, 2021

TABLE OF CONTENTS

	<i>Page</i>
PRELIMINARY STATEMENT	1
BACKGROUND	6
A. Class certification and Plaintiffs’ flawed theories	6
B. Goldman Sachs’ evaluation and promotion processes	7
C. The analyses performed by both sides’ experts show that the three challenged processes did not cause gender disparities	10
1. Plaintiffs’ expert, Dr. Farber, found <i>no</i> gender differences in the challenged processes for many groups of Goldman Sachs professionals	10
2. Goldman Sachs’ expert, Dr. Shaw, found <i>no</i> gender differences caused by the challenged processes, or found only differences that favored women	13
3. Plaintiffs mischaracterize Exhibit 68 to Dr. Shaw’s report	16
SUMMARY JUDGMENT STANDARD	18
ARGUMENT	18
I. Plaintiffs have not established a <i>prima facie</i> case of disparate impact under Title VII	18
A. Plaintiffs wrongly claim that their statistical evidence shows substantial disparities caused by the challenged processes	19
1. Dr. Farber’s models are unreliable, produce results that are disputed, and—even if taken at face value—would not prove disparate impact	20
a. Dr. Farber’s models are flawed and unreliable	22
b. Even accepting Dr. Farber’s models at face value, Dr. Farber fails to show that the challenged processes caused any purported gender disparities	29
2. Dr. Shaw does not “agree” that the challenged processes caused gender disparities	31

B.	Plaintiffs’ evidence as to 360 reviews and promotions fails to make out a <i>prima facie</i> case for the independent reason that Plaintiffs have no justification for their decision to lump together the discrete components of those processes	33
1.	Plaintiffs’ failure to analyze the separate steps of the 360 review and promotion processes precludes summary judgment in their favor and requires summary judgment for Goldman Sachs	33
2.	Dr. Shaw’s analysis of cross-ruffing according to statutory standards and the Certification Order cannot salvage plaintiffs’ summary judgment motion	35
II.	Plaintiffs have not established a <i>prima facie</i> case of disparate impact under the NYCHRL	38
	CONCLUSION	41

TABLE OF AUTHORITIES

	<i>Page(s)</i>
Cases	
<i>Abram v. United Parcel Serv. of Am., Inc.</i> , 200 F.R.D. 431 (E.D. Wis. 2001)	26
<i>Arnow v. Aeroflot Russian Airlines</i> , 980 F. Supp. 2d 477 (S.D.N.Y. 2013).....	38
<i>Assistant Deputy Wardens/Deputy Wardens Ass’n v. City of N.Y.</i> , 2019 WL 405119 (E.D.N.Y. Aug. 26, 2019).....	39
<i>Ass’n of Car Wash Owners Inc. v. City of N.Y.</i> , 911 F.3d 74 (2d Cir. 2018).....	18
<i>ATA Airlines, Inc. v. Fed. Exp. Corp.</i> , 665 F.3d 882 (7th Cir. 2011)	21
<i>Bolden v. Walsh Const. Co.</i> , 688 F.3d 893 (7th Cir. 2012)	25
<i>Brnovich v. Democratic Nat’l Comm.</i> , 141 S. Ct. 2321 (2021)	21
<i>Campbell v. Nat’l R.R. Passenger Corp.</i> , 311 F. Supp. 3d 281 (D.D.C. 2018).....	37
<i>Carpenter v. Boeing Co.</i> , 456 F.3d 1183 (10th Cir. 2006)	21
<i>Coates v. Johnson & Johnson</i> , 756 F.2d 524 (7th Cir. 1985)	24
<i>Coser v. Moore</i> , 739 F.2d 746 (2d Cir. 1984).....	26, 37
<i>Davis v. Cintas Corp.</i> , 717 F.3d 476 (6th Cir. 2013)	34, 38
<i>Easterling v. Conn.</i> , 783 F. Supp. 2d 323 (D. Conn. 2011).....	27, 28
<i>EEOC v. Freeman</i> , 961 F. Supp. 2d 783 (D. Md. 2013).....	35

<i>EEOC v. Joint Apprenticeship Comm. of Joint Indus. Bd. of Elec. Indus.</i> , 164 F.3d 89 (2d Cir. 1998).....	27
<i>Fitchett v. City of N. Y.</i> , 2021 WL 964972 (S.D.N.Y. Mar. 15, 2021)	38
<i>Florida v. United States</i> , 885 F. Supp. 2d 299 (D.D.C. 2012)	24
<i>Grant v. Metro. Gov’t of Nashville & Davidson Cty.</i> , 446 F. App’x 737 (6th Cir. 2011)	35, 38
<i>Green v. Town of E. Haven</i> , 952 F.3d 394 (2d Cir. 2020).....	24
<i>Gibson v. Am. Broad. Cos.</i> , 892 F.2d 1128 (2d Cir. 1989).....	18, 28
<i>Guippone v. BH S & B Holdings LLC</i> , 737 F.3d 221 (2d Cir. 2013).....	1, 2, 20
<i>Kassman v. KPMG LLP</i> , 416 F. Supp. 3d 252 (S.D.N.Y. 2018).....	26
<i>Kirkland v. N.Y. State Dep’t of Corr. Servs.</i> , 520 F.2d 420 (2d Cir. 1975).....	28, 29, 35
<i>Mandala v. NTT Data, Inc.</i> , 975 F.3d 202 (2d Cir. 2020).....	23, 24
<i>Morales v. Quintel Ent.</i> , 249 F.3d 115 (2d Cir. 2001).....	3
<i>Morgan v. United Parcel Serv. of Am., Inc.</i> , 380 F.3d 459 (8th Cir. 2004)	23
<i>NAACP v. N. Hudson Reg’l Fire & Rescue</i> , 742 F. Supp. 2d 501 (D.N.J. 2010)	27, 28
<i>Ottaviani v. State Univ. of N.Y. at New Paltz</i> , 875 F.2d 365 (2d Cir. 1989)	22
<i>Ricci v. DeStefano</i> , 557 U.S. 557 (2009)	19
<i>Richardson v. City of N.Y.</i> , 2021 WL 1910689 (S.D.N.Y. May 12, 2021)	3, 4, 29

<i>Robinson v. Metro-N. Commuter R.R. Co.</i> , 267 F.3d 147 (2d Cir. 2001).....	20
<i>Smith v. City of Jackson</i> , 544 U.S. 228, 241 (2005).....	34
<i>Smith v. Xerox Corp.</i> , 196 F.3d 358 (2d Cir. 1999).....	<i>passim</i>
<i>Sobel v. Yeshiva Univ.</i> , 839 F.2d 18 (2d Cir. 1988).....	22
<i>Teasdale v. City of N.Y.</i> , 2013 WL 5300699 (E.D.N.Y. Sept. 18, 2013)	5, 39
<i>United States v. City of N.Y.</i> , 637 F. Supp. 2d 77 (E.D.N.Y. 2009)	27, 28
<i>United States v. City of Warren</i> , 759 F. Supp. 355 (E.D. Mich. 1991).....	28
<i>Urbont v. Sony Music Ent.</i> , 831 F.3d 80 (2d Cir. 2016).....	18
<i>Valentino v. U.S. Postal Serv.</i> , 674 F.2d 56 (D.C. Cir. 1982)	26, 27
<i>Wado v. Xerox Corp.</i> , 991 F. Supp. 174 (W.D.N.Y. 1998).....	3, 23
<i>Walker v. City of N.Y.</i> , 2014 WL 1244778 (S.D.N.Y. Mar. 26, 2014)	29
<i>Wal-Mart Stores, Inc. v. Dukes</i> , 564 U.S. 338 (2011)	25, 26, 27
<i>Watson v. Fort Worth Bank & Tr.</i> , 487 U.S. 977 (1988)	20
<i>Weiss v. Nat'l Westminster Bank, PLC</i> , 993 F.3d 144 (2d Cir. 2021).....	3
Statutes	
42 U.S.C. § 2000e-2(k)(1)	<i>passim</i>
N.Y.C. Code § 8-107(17)(a)	<i>passim</i>

PRELIMINARY STATEMENT

Plaintiffs’ motion for summary judgment on the *prima facie* element of their disparate impact claims is a work of fiction. It distorts and mischaracterizes the parties’ expert statistical evidence to claim that the experts *agree* that the three challenged processes—360 reviews, manager quartiling, and cross-ruffing for promotions—caused a classwide disparate impact, when the opposite is true: Goldman Sachs’ expert, Dr. Kathryn Shaw, certainly never “agree[d] that the 360 Review and Manager Quartiling practices result in lower compensation for women” (Pls.’ Mot. 12), and, in fact, Plaintiffs presented ***no*** evidence that ***any*** of the three challenged processes caused disparities against women on a classwide basis. Plaintiffs obtained class certification based on their promise that the evidence would show that “[t]hroughout the entire [16-year] class period and across the [three] Class divisions, female Associates and Vice Presidents at Goldman have received lower scores” on 360 reviews and lower “quartile placements,” and were “promoted [to managing director] . . . at a statistically significantly lower rate” than their male counterparts. (Pls.’ Class Cert. Br. (ECF No. 247) at 12, 15 (emphasis added).) But all Plaintiffs’ own expert was able to conclude is that, at most, the three supposedly uniform challenged processes caused sporadic, not classwide, disparities. This not only bars summary judgment for Plaintiffs, but requires summary judgment for Goldman Sachs, as described in Goldman Sachs’ motion for summary judgment.

First, Plaintiffs’ “statistical proof” does not remotely “establish conclusively” that “the Challenged Employment Practices cause an adverse impact on the Class.” (Pls.’ Mot. 2-3.) As Plaintiffs appear to recognize, the statistical analysis proffered by their expert, Dr. Farber, cannot be a basis for granting summary judgment *to Plaintiffs* unless, among other things, it is undisputed that he used “appropriate model specifications” in his regression models. (*Id.* at 17); *see Guippono*

v. *BHS & B Holdings LLC*, 737 F.3d 221, 225 (2d Cir. 2013) (“Summary judgment is appropriate” only if “undisputed facts” show “moving party is entitled to judgment as a matter of law.”).

Here, however, Dr. Farber’s models are fundamentally flawed. Dr. Farber aggregated data across all three large and diverse Divisions at issue—the Investment Banking Division (“IBD”), Investment Management Division (“IMD”), and Securities—for 16 years despite unchallenged evidence that the independent and highly differentiated businesses within those Divisions apply the three challenged processes independently, and confirmation from basic statistical tests that pooling all this data into one model produces unreliable results.¹ Further, because Dr. Farber aggregated data across Divisions, he acknowledged that his model cannot “measure whether disparities for each of the challenged processes differ by [D]ivision” and so could not say if there were gender disparities in any particular Division. (GS 56.1 Resp. ¶ 23, GS Ex. 16 at 71:11-15, 85:17-86:4 (Farber Tr.).)

Dr. Farber also impermissibly “‘gerrymander[ed]’ [his] data to skew the results of statistical analysis in [Plaintiffs’] favor,” *Smith v. Xerox Corp.*, 196 F.3d 358, 369-70 (2d Cir. 1999), by completely excluding from his analysis associates and vice presidents in Goldman Sachs’ remaining, fourth revenue-generating Division (the Merchant Banking Division (“MBD”)), who were subject to the same challenged processes. And, lastly, Dr. Farber omitted from his models variables he concedes informed Goldman Sachs’ evaluation and promotion of its professionals, including obviously critical factors such as employee production. (*Infra* 23.) As a matter of law, flawed (and disputed) statistical analysis cannot compel a finding in Plaintiffs’ favor as required for Plaintiffs to prevail at summary judgment. To the contrary, courts have held that a

¹ See GS Decert Br. 13-16, 30-31, 36-37 (ECF No. 1224); GS 56.1 Resp. ¶ 23, Pls.’ Ex. 17 ¶¶ 177, 194, 218 (Shaw Rpt.).

plaintiff's expert's "fail[ure] to account" for employee "performance" is such a significant error that it renders the expert's statistical analysis incapable of even "creat[ing] a triable issue of fact on plaintiffs' disparate impact claims" when opposing a defendant's summary judgment motion. *Wado v. Xerox Corp.*, 991 F. Supp. 174, 184-85 (W.D.N.Y. 1998).²

Moreover, even if the Court accepts Dr. Farber's flawed models in considering Plaintiffs' motion for summary judgment (which would be improper), Dr. Farber's finding of a patchwork of disparities in only some parts of the class where supposedly uniform processes were applied compels a finding that the challenged processes themselves did not cause any purported classwide disparate impact. Specifically, Dr. Farber found that 360 reviews and manager quartiling produced no gender disparities for three full years of the class period (2016 to 2018), and that promotions favored women throughout the class period for five of 14 promotion cycles (2002, 2003, 2010, 2012, 2015). (GS 56.1 Resp. ¶¶ 58, 61, Pls.' Ex. 58 ¶ 117, App. A Table 28 (Farber Rebutl. Rpt.).) As explained in Goldman Sachs' motion for summary judgment, a "decision-making process either caused disparate impact or it did not," *Smith*, 196 F.3d at 369-70 (affirming summary judgment to defendant on a disparate impact claim), and if the three challenged employment processes are "the root cause of the inequality, their impact would presumably be felt and ascertainable across [all] [] job groups" and time periods where the processes were applied, which even Plaintiffs' expert found was not the case here. *Richardson v. City of N.Y.*, 2021 WL 1910689, at *9 (S.D.N.Y. May

² Although Goldman Sachs vigorously disputes the validity and reliability of Dr. Farber's flawed regression models (*see* GS *Daubert* Br. 36-55), Goldman Sachs does not rely on the flaws in Dr. Farber's models in support of its cross-motion for summary judgment, because Goldman Sachs is entitled to summary judgment on the basis of Dr. Farber's analysis even if accepted at face value. Thus, the disputed reliability of Dr. Farber's models is not a basis to deny Goldman Sachs' motion for summary judgment. *Weiss v. Nat'l Westminster Bank, PLC.*, 993 F.3d 144, 162 (2d Cir. 2021) ("immaterial" disputes about facts do not bar summary judgment); *see Morales v. Quintel Ent.*, 249 F.3d 115, 121 (2d Cir. 2001) (when parties file cross-motions for summary judgment, "each party's motion must be examined on its own merits").

12, 2021) (Oetken, J.) (denying class certification). Accordingly, Plaintiffs’ own statistical analysis demonstrates that Defendants, not Plaintiffs, are entitled to summary judgment on Plaintiffs’ disparate impact claims.

Plaintiffs try to manufacture a false agreement between the parties’ experts—supposedly in support of Plaintiffs—by misrepresenting the conclusions of Goldman Sachs’ statistical expert, labor economist Dr. Kathryn Shaw. Plaintiffs make the astounding (and unsupported) assertion that Dr. Shaw “concede[d] that there is a statistically significant gender pay gap caused by the Challenged Employment Practices.” (Pls.’ Mot. 3.) That is false. Dr. Shaw’s entire 248-page report is a point-by-point refutation of Dr. Farber’s methodologies and conclusions, and Dr. Shaw expressly concluded that “[t]he statistical evidence rejects the claim that there is a common, systemic gender gap with respect to the 360 review process, the assignment of manager quartiles, or promotion to [managing director] at GS during the relevant period,” and that, as a result, there was “no impact of such a gap on compensation.” (GS 56.1 Resp. ¶ 51, Pls.’ Ex. 17 ¶ 29 (Shaw Rpt.).) Plaintiffs are well aware that Dr. Shaw’s report conclusively debunks Dr. Farber’s analyses, along with Plaintiffs’ *prima facie* case of disparate impact.

Second, to challenge an entire “decisionmaking process” under Title VII, as Plaintiffs seek to do here, Plaintiffs must first “demonstrate to the court that the elements of a respondent’s decisionmaking process are not capable of separation for analysis.” 42 U.S.C. § 2000e-2(k)(1)(B)(i). Plaintiffs make no attempt—none—to meet this burden and offer no justification for their decision to challenge the *overall* 360 review, manager quartiling, and promotions processes without first meeting Title VII’s express requirement that Plaintiffs identify the specific “element[]” of each “particular employment practice,” and prove which particular element caused a disparate impact on women. *Id.* In fact, Plaintiffs concede that the promotions process can be

“split” into two “distinct pieces” (nominations and cross-ruffing) (Pls.’ Mot. 6, 23-24), and it is indisputable that “cross-ruffing” was the only process certified for class treatment (Class Cert. Order (ECF No. 578) at 28). Rather than confront their failure of proof on this indispensable aspect of their Title VII disparate impact claim, Plaintiffs pretend that there is no such requirement, citing outdated authorities from the *1970s*—nearly two decades before Congress codified 42 U.S.C. § 2000e-2(k)(1)(B) *in 1991*. Plaintiffs’ decision not to differentiate among the discrete elements of any of the three challenged processes alone is a sufficient basis to deny Plaintiffs’ motion, and to grant summary judgment to Goldman Sachs on Plaintiffs’ Title VII disparate impact claim.

Finally, Plaintiffs have failed to satisfy their *prima facie* case of disparate impact under the New York City Human Rights Law (“NYCHRL”). Plaintiffs incorrectly assert that they should be granted summary judgment on their NYCHRL claim “even if this Court finds that summary judgment is not appropriate for Plaintiffs’ Title VII *prima facie* case” because “the NYCHRL disparate impact standard is more lenient than Title VII.” (Pls.’ Mot. 24-25, 27.) Although the NYCHRL differs from Title VII in certain respects, Plaintiffs concede that it is the same in one crucial respect: Plaintiff must “prov[e] that the Challenged Employment Practices *cause* an adverse impact on the Class,” which they have failed to do. (Pls.’ Mot. 3 (emphasis added).) Tellingly, in the sole case Plaintiffs cite in support of their contention that they are entitled to summary judgment on their NYCHRL disparate impact claim, Judge Matsumoto granted summary judgment *to defendants* because the plaintiff’s statistical evidence was insufficient to establish disparate impact. *Teasdale v. City of N.Y.*, 2013 WL 5300699, at *9, 13 (E.D.N.Y. Sept. 18, 2013), *aff’d* 574 F. App’x 50 (2d Cir. 2014).

Moreover, as with their Title VII claim, Plaintiffs’ failure to analyze the certified cross-ruffing process *at all* is fatal to their NYCHRL claim. The Court only certified cross-ruffing for class treatment, which Dr. Farber never analyzed—he addressed overall promotions only. Plaintiffs’ failure of proof as to the certified process of cross-ruffing not only precludes summary judgment in their favor, it requires summary judgment for Goldman Sachs.

Plaintiffs’ motion for summary judgment should be denied.

BACKGROUND

A. Class certification and Plaintiffs’ flawed theories

In its March 30, 2018 Order, the Court certified a class consisting of female associates and vice presidents in revenue-producing roles within IBD, IMD, and Securities. (Class Cert. Order at 4, 49.) Under the Order, Plaintiffs may seek to prove that the three challenged processes (360 reviews, manager quartiling, and cross-ruffing) had a disparate impact on class members, and to prove a “disparate treatment claim predicated on the statistical evidence of disparate impact.” (*Id.* at 28, 41, 49.)³ In fact, Plaintiffs promised the Court that the evidence would show that “[t]hroughout the entire [16-year] class period and across the [three] Class divisions, female Associates and Vice Presidents at Goldman have received lower scores” on 360 reviews and lower “quartile placements,” and were “promoted [to managing director] . . . at a statistically significantly lower rate” than their male counterparts. (Pls.’ Class Cert. Br. at 12, 15.)

The Court did not certify promotions generally, but instead only “cross-ruffing.” *See, e.g.*, Oct. 7, 2019 Order (ECF No. 873) at 5 (The Court certified a class “with respect to three specific

³ As explained further in Defendants’ motion to decertify the class (ECF No. 1224, at 27 n.15), because Plaintiffs withdrew their motion to certify a Rule 23(b)(2) class seeking injunctive relief, Defendants have objected that the “phased” framework for trying Title VII class claims, which is premised on adjudicating classwide injunctive relief at Phase I, is not appropriate for trial of Plaintiffs’ backpay and punitive damages claims. Defendants preserve that objection here.

policies: 360 reviews, quartiling, and cross-ruffing.”); Class Cert. Order at 28 (“Whether Goldman Sachs discriminated against the class via its use of 360 reviews, quartiling, or cross-ruffing raises yes-or-no questions that can each be answered ‘in one stroke.’”). In response to an interrogatory that asked for “each employment practice or policy” that Plaintiffs would challenge, Plaintiffs identified only one promotion-related practice: “the firm wide ‘cross ruffing’ process for promotion from Vice President to Managing Director.” (GS 56.1 Resp. ¶ 13, GS Ex. 31 at 3-5 (Pls.’ 4/10/19 R&Os to Defs.’ Third Set of Interrogs.).)

Since class certification, the parties have engaged in three years of merits discovery based on Plaintiffs’ supposed theories of common proof and their promise that this case could be tried on a classwide basis, consistent with Rule 23, substantive law, and due process. In that time, further development of the record has consistently supported Goldman Sachs’ position that the challenged processes caused no disparate impact to class members.

B. Goldman Sachs’ evaluation and promotion processes.

During the class period, which begins July 7, 2002 (for New York-based class members) and September 10, 2004 (for class members outside of New York), Goldman Sachs was organized into four revenue-generating Divisions and numerous other non-revenue-generating Divisions. (*See* Class Cert. Order at 4.)

Class members are associates and vice presidents—two of the five designations of “corporate level” at Goldman Sachs (SAC (ECF No. 411) ¶¶ 61, 62),⁴ who work in revenue-generating roles, in only three of Goldman Sachs’ four revenue-generating Divisions—IBD, IMD,

⁴ The others are analyst, extended managing director, and participating managing director. (*See* GS Ex. 13 at 20-21, 30 (Dunleavy Rpt.).)

and Securities.⁵ Plaintiffs claim that the Firm’s 360 reviews, manager quartiling, and cross-ruffing processes disparately impacted female associates and vice presidents in IBD, IMD, and Securities (but not in the fourth revenue-producing division, MBD) during the class period. (Class Cert. Order at 4.)

A brief description of the three challenged processes as pertinent to this motion is as follows:

1. 360 Reviews. The 360 review process was a multi-step “framework[] for providing performance feedback” to Goldman Sachs professionals. (GS 56.1 Resp. ¶ 1, GS Ex. 18 ¶ 3 (Abramian-Katz Decl.); GS Ex. 22 ¶ 12 (Keefe Decl.).) The first step of the annual 360 review process was for Goldman Sachs professionals to select approximately “8 to 12 evaluators” who were “superiors, peers, and subordinates.” (Pls.’ Mot. 4.) Next, “[t]he scores across the review categories were combined to form an average review” (*id.* at 5), and a summary was prepared that incorporated the manager’s own evaluation of the professional.⁶ Last, the manager delivered the overall average 360 review score or, in later years, overall manager rating and a narrative feedback to the professional, typically in person.⁷

2. Manager Quartiling. Goldman Sachs used manager quartiling as a tool to evaluate relative “performance, contribution, and potential.” (Pls.’ Mot. 5.) The purpose of manager quartiling was to “[i]dentify top, middle, and bottom performers (relative to peers)” by distributing

⁵ This brief refers to Goldman Sachs’ Divisions and structure as they existed in November 2018. Plaintiffs have adduced no evidence that postdates that period.

⁶ GS 56.1 Resp. ¶ 6, Pls.’ Ex. 2 at 280:5-11 (Heller Sberlati Tr.); Pls.’ Ex. 1 at 280:1-18 (Kung Tr.); GS Ex. 9 ¶¶ 38, 65 (Cascio Rpt.); Pls.’ Ex. 17 ¶¶ 98-99 (Shaw Rpt.); GS Ex. 11 at 5 (4/19/21 Cascio Rebt. Rpt.). Through 2009, 360 reviewers scored each competency on a 1-5 scale; this was changed to a 1-9 scale in 2010; and after 2016 three ratings were used. (*See* Pls.’ 56.1 ¶ 5.)

⁷ GS 56.1 Resp. ¶ 6, Pls.’ Ex. 2 at 279:15-23 (Heller Sberlati Tr.); Pls.’ Ex. 1 at 290:6-12 (Kung Tr.); Pls.’ Ex. 17 ¶¶ 98-99 (Shaw Rpt.).

professionals into four or five manager “quartiles.”⁸ Managers used their discretion in making these determinations, taking into account many factors including the importance of the unique role each professional played in their business unit, the manager’s view of their performance, and the future needs of the business.⁹

3. Cross-Ruffing. The promotion process from vice president to managing director at Goldman Sachs involved two steps: (i) a “nomination” process; and (ii) a vetting process known as “cross-ruffing” for those candidates who were nominated for promotion. (Pls.’ Mot. 6, 23-24; *see also* GS Ex. 16 at 183:17-184:11 (Farber Tr.).)¹⁰ At the “[f]irst” step (nominations), managing directors, “business unit heads,” and Division heads helped to “create a division-wide list of candidates” who would be selected for the “cross-ruffing” vetting process.¹¹ Plaintiffs have identified no classwide centralized process for nominations that applied throughout the class period. Indeed, from 2002-2007, any managing director or participating managing director could nominate any vice president for consideration. (GS 56.1 Resp. ¶ 13, Pls.’ Ex. 39 at -564 (GS0113548); Pls.’ Ex. 53 at -973 (GS0215973); Pls.’ Ex. 54 at -972 (GS0164972).)

At the “[s]econd” step (cross-ruffing), each member of each Division’s “cross-ruffing committee” was “assigned a list” of nominated candidates to evaluate by “interview[ing] Managing Directors who [were] knowledgeable about the candidate”; each then “prepar[ed] a one-

⁸ GS 56.1 Resp. ¶ 8, GS Ex. 4 at -659 (GS0375658); Pls.’ Ex. 18 ¶ 35 (Farber Rpt.); *see also* Pls.’ 56.1 ¶ 8; Pls.’ Ex. 1 at 29:6-19, 316:15-317:2, 317:7-24 (Kung Tr.); Pls.’ Ex. 3 at 182:16-183:22 (Larson Tr.); Pls.’ Ex. 4 at 9:12-10:13 (Landman II Tr.).

⁹ GS 56.1 Resp. ¶¶ 9-11, GS Ex. 20 ¶¶ 11, 13-14 (Giovanni Decl.); Pls.’ Ex. 21 at -290 (GS0153290), Pls.’ Ex. 22 at -355 360 (GS0109353).

¹⁰ Plaintiffs only sought certification of “the ‘cross-ruffing’ promotion process.” (Pls.’ Class Cert. Reply (ECF No. 310) at 7-8.)

¹¹ Pls.’ Mot. 6; *see also* Pls.’ Ex. 17 ¶¶ 118-20 (Shaw Rpt.); Pls.’ Ex. 2 at 232:22-233:9 (Heller Sberlati Tr.); Pls.’ Ex. 3 at 237:17-25 (Larson Tr.).

page summary review sheet.” (Pls.’ Mot. 7.) Following the interviews, the divisional cross-ruffing team “rank[ed] the list of candidates in order of preference for promotion,” and “[t]he Division Heads then review[ed] this list” and “create[d] their own ranked list.” (*Id.*)

It is undisputed that managers across the three Divisions implemented these three processes independently and differently. (*See* GS 56.1 Resp. ¶ 23, Pls.’ Ex. 17 ¶¶ 32, 89-135, 186 (Shaw Rpt.), ECF Nos. 1225-1, 1225-2, 1225-3 (compilations from declarations of dozens of business leaders explaining how they uniquely apply the challenged processes).) Plaintiffs contend that none of the processes changed in any ways relevant to their claims over the class period. (*See* Pls.’ 56.1 ¶¶ 1-22.)

C. The analyses performed by both sides’ experts show that the three challenged processes did not cause gender disparities.

Plaintiffs’ primary contention in their motion for summary judgment on their *prima facie* case of disparate impact is that it is undisputed that “the Challenged Employment Practices cause *an adverse impact on the Class*.” (Pls.’ Mot. 3 (emphasis added).) Plaintiffs’ statements describing the parties’ experts’ analyses are not accurate.

1. Plaintiffs’ expert, Dr. Farber, found *no* gender differences in the challenged processes for many groups of Goldman Sachs professionals.

Plaintiffs’ statistical expert, Dr. Farber, built multiple regression models to analyze “[w]hether there is statistical evidence of a difference in how men and women fare in Goldman’s performance review system” and “[w]hether there is statistical evidence of disparities between men and women in promotion rates from Vice President positions to positions as Managing Directors.” (Pls.’ Ex. 18 ¶ 6 (Farber Rpt.).)

Plaintiffs contend that it is undisputed that Dr. Farber employed “generally accepted statistical method[s]” in constructing his regression models, and used “appropriate model specifications” to “control for relevant worker characteristics in the data that may affect the

outcome” of the three challenged processes. (Pls.’ Mot. 8, 17.) That is false. Dr. Farber’s models have numerous, significant flaws. Dr. Farber improperly aggregated data across all three class-relevant Divisions for all women and men in all class-relevant roles for 16 years (from 2003 to 2018),¹² even though the three challenged processes are implemented differently across Divisions (*supra* n.1). As noted, because Dr. Farber aggregated data across all three Divisions, he cannot say if there were gender disparities in any particular Division. (GS 56.1 Resp. ¶ 23, GS Ex. 16 at 71:11-15, 85:17-86:4 (Farber Tr.); *supra* 2.) It is therefore false to state, as Plaintiffs have, that Dr. Farber found gender disparities across the class. In addition, Dr. Farber omitted entirely from his models what is perhaps the most important factor to consider in assessing the application of “performance review” processes (Pls.’ Mot. 4): employee production metrics.¹³ Dr. Farber acknowledged the significance of employee production in one of his reports in this case: “Economists define sex discrimination in pay to be differences in pay between men and women that cannot be explained by differences in *productivity-related* characteristics.” (GS 56.1 Resp. ¶ 23, GS Ex. 15 ¶ 33 (Oct. 30, 2013 Farber Rpt.) (emphasis added).). Dr. Farber’s failure here to include any controls for production metrics in his models renders his results unreliable.

As to 360 reviews, Plaintiffs contend that “Dr. Farber found statistical evidence of adverse impact against Class Members in the 360 Review from 2002 to 2015.” (Pls.’ Mot. 8.) This statement is not correct. Dr. Farber performed no analysis of 360 reviews in 2002 because the firmwide 360 review process did not exist until 2003, when it was only partially implemented, leaving no sufficient data to analyze across three Divisions.¹⁴ Dr. Farber nevertheless purported

¹² GS 56.1 Resp. ¶¶ 23, Pls.’ Ex. 18 ¶ 77, Tables 7, 8, 11, 12, 19, 20, 28 (Farber Rpt.); GS Ex. 16 at 71:3-10, 72:5-73:6 (Farber Tr.).

¹³ GS 56.1 Resp. ¶¶ 23, 25, 28, 30-40, Pls.’ Ex. 17 ¶¶ 149, 193, 214 & Exs. 21, 24 (Shaw Rpt.).

¹⁴ GS 56.1 Resp. ¶¶ 1, 5, 25, GS Ex. 7 at 38:14-19 (9/5/13 Landman Tr.). Dr. Farber only claimed that data regarding 360 scores was available for 2003 to 2015, thus admitting that no data was

to reach conclusions about 2003 and forward (not 2002), segmenting his analysis into three time periods (2003 to 2009, 2010 to 2015, and 2016 to 2018). (Pls.’ Ex. 18 Tables 11, 12, 19, 20 (Farber Rpt.).) In presenting Dr. Farber’s findings only through 2015 (Pls.’ Mot. 8), Plaintiffs omit that Dr. Farber found “no statistically significant differences between male and female” associates or vice presidents in their 360 review scores *after* 2015. (GS 56.1 Resp. ¶¶ 25, 26, 33, Pls.’ Ex. 18 ¶ 114 & Tables 19, 20 (Farber Rpt.); Pls.’ Ex. 17, Ex. 69 (Shaw Rpt.) (summarizing Dr. Farber’s findings).) Plaintiffs also leave out the fact that Dr. Farber did not separately analyze *any* of the individual elements that make up the 360 review cycle, such as reviewer selection; individual peer, subordinate, and manager reviews; the averaging process by which individual reviews were converted into overall 360 scores or ratings; or the impact of the different skills and competencies assessed in 360 reviews. (GS 56.1 Resp. ¶ 23, Pls.’ Ex. 18 ¶¶ 86-88 & Tables 11, 12 (Farber Rpt.).) Dr. Farber’s failure to identify which element of the 360 review process purportedly caused gender disparities renders his analysis useless.

As to manager quartiling, Plaintiffs contend that “Dr. Farber found statistical evidence of adverse impact against Class Members in the Manager Quartile from 2002 to 2015.” (Pls.’ Mot. 9.) Plaintiffs again omit any mention of Dr. Farber’s express finding of “no statistically significant differences between male and female” associates or vice presidents in manager quartiling *after* 2015. (GS 56.1 Resp. ¶ 33, Pls.’ Ex. 18 ¶ 114 & Tables 19, 20 (Farber Rpt.).) Plaintiffs also ignore that Dr. Farber’s aggregate models cannot detect if disparities exist in any particular Division.

available for 2002. (GS 56.1 Resp. ¶¶ 1, 5, 25, Pls.’ Ex. 18 ¶¶ 10, 45 (Farber Rpt.).) Dr. Farber’s analysis of the 360 review process for IBD and IMD is based on only 17 observations for the entirety of 2003 and 2004. (GS 56.1 Resp. ¶ 25, GS Ex. 17 (excerpt of Farber backup).)

With respect to the certified “cross-ruffing process” (Class Cert. Order at 28), Plaintiffs claim that Dr. Farber’s “statistical evidence also shows that Goldman’s cross-ruffing process has disadvantaged female Vice Presidents throughout the Class Period.” (Pls.’ Mot. 13.) This statement is also untrue. In reality, Dr. Farber admitted that he “*did not do any analysis of cross-ruffing*” and instead chose to analyze “the overall effect” of the “promotion process, not cross-ruffing specifically.” (GS 56.1 Resp. ¶ 58, GS Ex. 16 at 137:9-10, 149:10-16 (Farber Tr.) (emphasis added).) The result of Dr. Farber’s choice is that the Plaintiffs have not produced *any* statistical evidence that cross-ruffing caused any disparate impact. Although he acknowledged that it was possible to “disaggregat[e] the promotions process into separate steps,” Dr. Farber did not do so. (GS 56.1 Resp. ¶¶ 58, 63, Pls.’ Ex. 58 ¶ 81 (Farber Rebutl. Rpt.).) He thus confirmed that he was “not offering any opinion about disparities between rates of promotion of men and women who are on cross-ruffing lists.” (GS Ex. 16 at 185:18-22 (Farber Tr.).) Thus, irrefutably, Plaintiffs have not produced any evidence in support of their certified cross-ruffing claim.

Nor did Dr. Farber find that overall promotions processes “disadvantage[] female Vice Presidents throughout the Class Period.” (Pls.’ Mot. 13.) In reality, Dr. Farber found no disparities, or disparities that favored women, in over a third (five out of 14) of the promotion cycles during the class period. (GS 56.1 Resp. ¶ 58, 61, Pls.’ Ex. 58 ¶ 117, App. A, Table 28 (Farber Rebutl. Rpt.).)

2. Goldman Sachs’ expert, Dr. Shaw, found *no* gender differences caused by the challenged processes, or found only differences that favored women.

Goldman Sachs’ expert, Dr. Shaw, conducted regression analyses of the same time periods and title groupings that Dr. Farber used, but she ran separate regression models for each Division (IMD, IBD, and Securities). (GS 56.1 Resp. ¶¶ 25, 33, Pls.’ Ex. 17, ¶ 187, Exs. 21, 24, 27-29 (Shaw Rpt.).) Dr. Farber testified that there was sufficient data to run separate models in this way.

(GS 56.1 Resp. ¶ 23, GS Ex. 16 at 176:21-177:2 (Farber Tr.) (“Q. Did you ever determine that the sample sizes in the divisions were too small to run your model separately by division? A. No, I don’t think that was an issue.”).)

Dr. Shaw also ran her models controlling for the variables Dr. Farber has acknowledged are important to include in any regression model but mysteriously omitted in his studies here—namely measures of productivity that “can be used to help delineate the different skills that are valued across jobs within” Goldman Sachs. (GS 56.1 Resp. ¶¶ 23-26, 28, 33, Pls.’ Ex. 17 ¶ 74 (Shaw Rpt.).) For example:

- In IBD, data on production included “total revenue generated across all projects within a given year.” (GS 56.1 Resp. ¶ 25, Pls.’ Ex. 17 ¶ 74 (Shaw Rpt.).)
- In IMD, data included “information on ‘Alpha and Risk’ and ‘Investment Culture’ scores for professionals,” which “measure[d] a professional’s contributions to the performance and risk management of investment portfolios” and “the manager’s evaluation of qualitative aspects of the professional’s performance.” (*Id.*)
- In Securities, data included “the actual financial contribution of each professional within a given year,” defined depending on the professional’s exact function. (*Id.*)

Managers at Goldman Sachs have testified to the commonsense point that performance and financial contribution reflected in this data was a key component that factored into all three challenged processes, to different degrees depending on the manager and specific role.¹⁵

Plaintiffs contend that Dr. Shaw “concede[d] that there is a statistically significant gender

¹⁵ See GS 56.1 Resp. ¶ 23, GS Ex. 29 ¶ 7 (Russell Decl.) (“commercial production” influenced 360 reviews); GS Ex. 19 ¶ 12 (S. Cohen Decl.) (“commercial impact” influenced 360 reviews); GS Ex. 18 ¶ 12 (Abramian-Katz Decl.) (“profit and loss for traders . . . and revenue generation” in certain Business Units influenced quartiling); GS Ex. 21 ¶ 12 (Goddeeris Decl.) (“objective performance metrics, including profits and losses” influenced quartiling for Securities traders); GS Ex. 25 ¶ 10 (Lopez Decl.) (same for electronic sales trading and sales acquisition); GS Ex. 24 ¶ 26 (Kozlowski Decl.) (“commercial productivity” accounted for in cross-ruffing); GS Ex. 30 Ex. 1 ¶ 11 (Scher Decl.) (same for “productivity”); GS Ex. 26 ¶ 19 (Mora Decl.) (same).

pay gap caused by the Challenged Employment Practices.” (Pls.’ Mot. 3.) That also is not true.

As Dr. Shaw stated in her report, her opinion is the exact opposite:

The statistical evidence *rejects the claim* that there is a common, systemic gender gap with respect to the 360 review process, the assignment of manager quartiles, or promotion to [managing director] at GS during the relevant period.

(GS 56.1 Resp. ¶ 51, Pls.’ Ex. 17 ¶ 29 (Shaw Rpt.); *see also id.* ¶¶ 185-239 (Dr. Shaw’s analysis).)

As a result, Dr. Shaw concluded that “[t]he statistical evidence rejects the claim that alleged gender gaps in the challenged processes *cause* an alleged gender gap in compensation.” (*Id.* ¶¶ 29, 240.)

With respect to 360 reviews and manager quartiling specifically, Dr. Shaw found that there were only a few, sporadic gender gaps when she ran separate regression models for each Division, as shown below:

Table 1

	Models	IBD		IMD		Securities	
		Junior Bankers	Senior Bankers	Associates	Vice Presidents	Associates	Vice Presidents
Manager quartiles	pre-2016	⊖	⊖	⊖	⊖	⊖	⊖
	2016 to 2018	⊖	⊖	⊖	⊖	⊖	⊖
360 Review (OMR)	pre-2010	⊖	⊖	⊖	⊖	⊖	⊖
	2010 to 2015	⊖	⊖	⊖	⊖	⊖	⊖
	2016 to 2018	⊖	⊖	⊖	⊖	⊖	⊖

(GS 56.1 Resp. ¶¶ 25, 33, Pls.’ Ex. 17 ¶¶ 193, 214 & Exs. 21, 24 (Shaw Rpt.).) Dr. Shaw concluded that her analysis “rejects Plaintiffs’ claims of systemic bias.” (*Id.* ¶ 33.)

As to promotions, Dr. Shaw analyzed separately the “two key steps in the promotion process: (i) selection to the nominee list, and (ii) cross-ruffing,” and additionally analyzed the promotions process as a whole. (GS 56.1 Resp. ¶¶ 13, 58, 63, Pls.’ Ex. 17 ¶¶ 232, 237, Exs. 27-

29 (Shaw Rpt.).)¹⁶ As shown below, Dr. Shaw found that “women [were] actually nominated at a higher rate than men in two of the three divisions” (IMD and Securities); that once nominated, “women nominees [were] promoted no less often than men” or were “more likely to be promoted than men” in all three Divisions; and that when the overall promotions process is analyzed, “women VPs [were] no less likely to be promoted to [managing director] than their male counterparts are.” (GS 56.1 Resp. ¶¶ 58, 61, Pls.’ Ex. 17 ¶¶ 235-37, Ex. 27-29 (Shaw Rpt.).)

Table 2

	IBD	IMD	Securities
Nominations	=	+	+
Promotions (among those nominated)	=	=	=
Overall Promotion Process	=	=	+

As Dr. Shaw explained, “[p]romotion to [managing director] is a significant business decision for GS,” and “[t]he fact that women fare as well as or better than men in the promotion process is strongly inconsistent with the claim that GS is biased against women.” (GS 56.1 Resp. ¶ 58, Pls.’ Ex. 17 ¶ 239 (Shaw Rpt.).)

3. Plaintiffs mischaracterize Exhibit 68 to Dr. Shaw’s report.

Plaintiffs mischaracterize Dr. Shaw’s opinions in other ways too. Plaintiffs contend that Exhibit 68 to Dr. Shaw’s report reflects her agreement “both that there is a statistically significant gender pay gap at Goldman Sachs, and that the 360 Review and Manager Quartile are a major cause of it.” (Pls.’ Mot. 12.) Neither claim is true.

¹⁶ Dr. Shaw explained that while “nomination is not itself a challenged process in this matter,” she analyzed it in order to respond to Dr. Farber’s analysis of “promotions as a whole,” which are comprised of nominations and cross-ruffing. (GS 56.1 Resp. ¶ 58, Pls.’ Ex. 17 ¶ 232 (Shaw Rpt.).)

As to a supposed “gender pay gap,” the Court did not certify claims regarding compensation untethered to the three challenged processes,¹⁷ yet Dr. Farber conducted irrelevant “analyses of compensation at Goldman Sachs unrelated to the impact of the three challenged processes.” (GS 56.1 Resp. ¶ 51, GS Ex. 16 at 182:6-11 (Farber Tr.)) In response, Dr. Shaw “evaluate[d] these analyses” and presented her analysis of Dr. Farber’s results in Exhibit 68, concluding the exact opposite of what Plaintiffs claim: that there was no “evidence of a common gender gap” in compensation. (GS 56.1 Resp. ¶¶ 51-52, Pls.’ Ex. 17 ¶ 248, Ex. 68 (Shaw Rpt.)) Plaintiffs mischaracterize Dr. Shaw’s opinion by pointing to “Rows 4 & 9” of Exhibit 68, which are interim calculations that are *not* Dr. Shaw’s full analysis, and claiming falsely that they include “every single control variable that she advocates for.” (Pls.’ Mot. 12 n.53.) In fact, when all of Dr. Shaw’s control variables are included (as shown in the very next rows of Exhibit 68), there is no classwide gender pay gap—a point that is underscored by Dr. Shaw’s alternate compensation analysis, a median regression, which, as shown below, demonstrates that *there are no significant gender gaps in compensation for any group of class-relevant employees*:

Table 3

	IBD		IMD		Securities	
	Junior Bankers	Senior Bankers	Associates	Vice Presidents	Associates	Vice Presidents
Median Regression	=	=	=	=	=	=

(GS 56.1 Resp. ¶¶ 46, 51-52, Pls.’ Ex. 17 Ex. 35 (Shaw Rpt.))

¹⁷ Both this Court and Judge Lehrburger have held repeatedly that “Phase I will be devoted to generalized proof of the alleged discriminatory impact and treatment of women through three practices: 360 reviews, quartiling, and cross-ruffing,” and that “compensation may be a topic addressed in Phase I *insofar as it is connected to one or more of the three challenged processes*.” (June 25, 2019 Order (ECF No. 767) at 1-2 (emphasis added).)

Plaintiffs’ second claim about Exhibit 68, that it reflects Dr. Shaw’s “agree[ment] that the 360 Review and Manager Quartiling practices result in lower compensation for women” (Pls.’ Mot. 12), is equally untrue. Plaintiffs base this false assertion on the fact that 360 reviews and manager quartiles are factors that bear on compensation (*id.*)—an unsurprising result given that they are key performance metrics. But that has nothing to do with whether those processes *cause* a gender gap in compensation, which could only be true if Dr. Shaw’s analysis showed that 360 reviews and manager quartiles themselves disadvantage women. As already stated, Dr. Shaw reached the opposite conclusion, opining that “[t]he statistical evidence rejects the claim that alleged gender gaps in the challenged processes cause an alleged gender gap in compensation.” (GS 56.1 Resp. ¶ 51, Pls.’ Ex. 17 ¶ 29 (Shaw Rpt.).)

SUMMARY JUDGMENT STANDARD

“Summary judgment is appropriate only where ‘there is no genuine dispute as to any material fact and the movant is entitled to judgment as a matter of law.’” *Ass’n of Car Wash Owners Inc. v. City of N.Y.*, 911 F.3d 74, 80 (2d Cir. 2018). On Plaintiffs’ motion for summary judgment, the “court’s function is not to weigh the evidence, make credibility determinations or resolve issues of fact, but rather to determine whether, drawing all reasonable inferences from the evidence presented in favor of [Defendants], a fair-minded jury could find in [Defendants’] favor.” *Urbont v. Sony Music Ent.*, 831 F.3d 80, 88 (2d Cir. 2016). “The burden of demonstrating the lack of any genuine unresolved issues of fact rests on the moving party,” here, Plaintiffs. *Gibson v. Am. Broad. Cos.*, 892 F.2d 1128, 1132 (2d Cir. 1989).

ARGUMENT

I. Plaintiffs have not established a *prima facie* case of disparate impact under Title VII.

To establish a *prima facie* case of disparate impact under Title VII’s burden-shifting framework, Plaintiffs must prove that Goldman Sachs “uses ‘a particular employment practice that

causes a disparate impact on the basis of . . . sex.” *Ricci v. DeStefano*, 557 U.S. 557, 578 (2009). Plaintiffs have not done so, because Plaintiffs cannot prove that 360 reviews, manager quartiling, and cross-ruffing caused classwide disparities. Plaintiffs also cannot prove (i) disparate impact for any element of the 360 review or promotions processes, or (ii) show that those elements “are not capable of separation for analysis.” 42 U.S.C. § 2000e-2(k)(1)(B)(i). This failure of proof means not only that the Court should deny Plaintiffs’ motion for summary judgment, but also that summary judgment should be entered in favor of Goldman Sachs.

A. Plaintiffs wrongly claim that their statistical evidence shows substantial disparities caused by the challenged processes.

Plaintiffs’ “statistical proof” does not come close to “establish[ing] conclusively” that “the Challenged Employment Practices cause an adverse impact on the Class.” (Pls.’ Mot. 2-3.) To satisfy their burden, Plaintiffs must “present statistical evidence ‘of a kind and degree sufficient to show that the practice in question has caused [a disparate impact on women] because of their membership in a protected group.’” *Smith*, 196 F.3d at 365. They have not done so. As Defendants’ showed in their *Daubert* motion (ECF No. 1188 at 36-55), Dr. Farber’s statistical analyses are unreliable and, even if his opinions satisfied the threshold for admissibility (they do not), reasonable jurors would have ample reason to reject Dr. Farber’s single, pooled regression models because of his methodological flaws. Dr. Farber both inappropriately pooled data across all three Divisions and failed to account for indisputably significant variables, including employee production.

Beyond that, even if the Court were to accept Dr. Farber’s analyses at face value for purposes of Plaintiffs’ motion (and the Court should not), Dr. Farber found that the three challenged processes produced no gender disparities at all for large groups of class members during multiple years of the 16-year class period. (*See* Pls.’ Ex. 58 App. A, Table 28 (Farber Rebtl.

Rpt.); *see also* GS 56.1 Resp. ¶¶ 25, 33, 58, 61.) As explained in Goldman Sachs’ summary judgment motion, this inconsistent, sporadic patchwork of gender differences associated with the challenged processes, as reflected in Plaintiffs’ own expert report, eviscerates any notion that the challenged processes themselves were the cause of any purported disparities. Plaintiffs try to shore up that fatal shortcoming in Dr. Farber’s analysis by falsely stating that Dr. Shaw “concedes that there is a statistically significant gender pay gap caused by the Challenged Employment Practices,” when her opinions are the very opposite. (Pls.’ Mot. 3.) Of course, false characterizations are not a ground for summary judgment, which can be granted only on the basis of “undisputed facts.” *Guippone*, 737 F.3d at 225.

1. Dr. Farber’s models are unreliable, produce results that are disputed, and—even if taken at face value—would not prove disparate impact.

To make out a *prima facie* case of disparate impact, Plaintiffs’ statistical evidence must (i) show a “‘substantial’ or ‘significant’” gender “disparity,” and (ii) “be of a kind and degree sufficient to reveal a causal relationship between the challenged practice and the disparity.” *Robinson v. Metro-N. Commuter R.R. Co.*, 267 F.3d 147, 160 (2d Cir. 2001). No “rigid mathematical formula” determines whether statistics raise an inference of causation, *Watson v. Fort Worth Bank & Tr.*, 487 U.S. 977, 995 (1988), and “the substantiality of a disparity is judged on a case-by-case basis,” *Smith*, 196 F.3d at 366.

Despite the Supreme Court’s admonition that Plaintiffs’ burden to establish a *prima facie* case of disparate impact cannot be reduced to a “rigid mathematical formula,” *Watson*, 487 U.S. at 995, Plaintiffs contend that they are entitled to summary judgment merely because Dr. Farber’s aggregate statistics purport to find an average gender disparity across the three large and diverse Divisions that is “statistically significant” by more than “two standard deviations.” (Pls.’ Mot. 17-

19, 22-23.) This contention fundamentally misconstrues Plaintiffs’ burden at summary judgment and the meaning of statistical significance.

Simply presenting a regression model that spits out a statistically significant result does not establish causation. As the Supreme Court recently emphasized, “[s]tatistical significance may provide ‘evidence that something besides random error is at work,’ . . . but it does not necessarily determine causes.” *Brnovich v. Democratic Nat’l Comm.*, 141 S. Ct. 2321, 2343 n.17 (2021). Thus, even a “large number of standard deviations tells us nothing about what that ‘something’ [causing the gender disparity] is.” *Carpenter v. Boeing Co.*, 456 F.3d 1183, 1202 (10th Cir. 2006). This is because Dr. Farber’s regression models, like all regression models, have a significant limitation. As Dr. Farber explained at the class-certification hearing, regression models do not identify the cause of a gender disparity, but instead attempt to rule out *other causes* of gender disparities (like differences in education or performance), such that what remains (called the “residual”) “is an unexplained pay gap.” (ECF No. 345 at 277:5-14.) Put differently, the output of a regression model (*i.e.*, the “residual”) is what is “left over” and “not explained” by the factors included in the model. *ATA Airlines, Inc. v. Fed. Exp. Corp.*, 665 F.3d 882, 890 (7th Cir. 2011).¹⁸ Thus, as Dr. Farber himself admits, his regression models establish only that the gender disparities he purports to find are caused by one or more of the factors that are ***not in his model***, which Dr. Farber calls the “unmeasured factors,” and which, significantly, include objective measures of

¹⁸ More technically, when a regression is performed, a “coefficient” is derived for each factor (like education or prior work experience), which represents the relative impact of that factor on the outcome of the process being studied—here, 360 scores, manager quartile ratings, and promotions. (Pls.’ Ex. 18 ¶ 55 (Farber Rpt.).) For each individual employee, multiplying the factors included in the model by their coefficients predicts an outcome such as a 360 score. The “residual” is the difference between the actual score and the one predicted by the model—and so reflects the influence of all factors that are not in the model, which Dr. Farber calls the “unmeasured factors.” (*Id.*)

employees’ performance, such as P&L and other production measures—none of which Dr. Farber chose to include in his models despite his statements about the importance of taking such factors into consideration. (GS 56.1 Resp. ¶¶ 25, 33, Pls.’ Ex. 18 ¶¶ 54-55 (Farber Rpt.), GS Ex. 16 at 204:16-21, 233:21-234:14 (Farber Tr.); *see infra* n.19.)

Because the residual constitutes nothing more than the effect of factors left out of the regression model, one can infer based on process of elimination that the “residual” “represents the unlawful effect” of gender discrimination *only if* the regression model includes “all of the possible ‘legitimate’ (*i.e.*, nondiscriminatory) factors that are likely to significantly affect the dependent variable [here, 360 review scores, quartile ratings and promotions] and which could account for disparities in the treatment of male and female employees.” *Ottaviani v. State Univ. of N.Y. at New Paltz*, 875 F.2d 365, 367 (2d Cir. 1989). Plaintiffs acknowledge as much, recognizing that they cannot prevail on their summary judgment motion unless, among other things, it is undisputed that Dr. Farber used “appropriate model specifications” in his regression models. (Pls.’ Mot. 17); *see also Sobel v. Yeshiva Univ.*, 839 F.2d 18, 22 (2d Cir. 1988) (“As with any multiple regression analysis, the validity of the influence attributed to a particular variable will depend heavily on how accurately the model mimics the actual factors influencing the dependent variable, salary.”). Plaintiffs cannot meet that burden.

a. Dr. Farber’s models are flawed and unreliable.

Plaintiffs cannot meet their summary judgment burden first and foremost because Dr. Farber’s regression models are fatally flawed. Goldman Sachs catalogued those deficiencies in detail in its *Daubert* motion and summary judgment motion, and will not repeat all of those arguments here. But, at a minimum, even if his models were admissible, any one of the following flaws in Dr. Farber’s models would give a reasonable juror grounds to conclude that Dr. Farber’s models do not give rise to an inference of causation, barring summary judgment for Plaintiffs:

- Dr. Farber ignored employee productivity in his models, which he conceded was necessary to consider in studying the effect of gender on a company's evaluation and promotion processes.¹⁹ (GS 56.1 Resp. ¶ 23; *see* GS *Daubert* Br. 50-51; GS *Daubert* Opp. (ECF No. 1259) at 9-15 (discussing the extensive record evidence showing that, unsurprisingly, production metrics directly impact performance assessment at Goldman Sachs).) Dr. Farber's failure to account for employee production metrics, which are objective measures of performance, is fatal to his analysis. *See Morgan v. United Parcel Serv. of Am., Inc.*, 380 F.3d 459, 468 (8th Cir. 2004) ("A reasonable inference of such discrimination does not arise in this case" because "Plaintiffs' regressions fail to take into account past pay, *performance*, or both.") (emphasis added); *Wado*, 991 F. Supp. at 184 (plaintiffs' regressions do "not suffice to create a triable issue" because expert "failed to account for possible nondiscriminatory reasons for the disparities that he found," including "*performance* and skills ratings") (emphasis added).
- Dr. Farber inappropriately aggregated data from all three Divisions (IBD, IMD and Securities), even though (i) a Chow test, a standard and recognized statistical technique, established that pooling across the three large and diverse Divisions is not appropriate, (ii) Dr. Farber admitted that the three Divisions each had sufficient populations to analyze in a statistically valid way, and (iii) there is overwhelming evidence that managers across Goldman Sachs implement the three challenged processes independently and differently.²⁰ As a matter of law, a statistical analysis that wrongly "assum[es] 'that what is true for the whole is true for a[ll] subset[s] of the whole'" is a "fatal" defect even at the pleading stage, and is certainly no basis for a compelled finding of disparate impact. *Mandala v. NTT Data, Inc.*, 975 F.3d 202, 211 (2d Cir.

¹⁹ "Economists define sex discrimination in pay to be differences in pay between men and women that cannot be explained by differences in *productivity-related* characteristics." (GS 56.1 Resp. ¶ 23, GS Ex. 15 ¶ 33 (Oct. 30, 2013 Farber Rpt.) (emphasis added).).

²⁰ GS 56.1 Resp. ¶ 23, GS Ex. 16 at 176:21-177:2, 178:21-180:7 (Farber Tr.); *supra* 10; *see also* GS *Daubert* Br. 38-44.

2020) (affirming dismissal of discrimination claims based on flawed aggregated statistical evidence).²¹

- Dr. Farber “‘gerrymander[ed]’ [his] data to skew the results of statistical analysis in [Plaintiffs’] favor,” *Smith*, 196 F.3d at 369-70, by excluding from his analysis associates and vice presidents in Goldman Sachs’ remaining, fourth revenue-generating Division (MBD), who were subject to the same challenged processes. (*See* GS SJ Br. (ECF No. 1240) 4, 7, 26.)

Tellingly, when Dr. Farber’s errors are corrected with respect to the three Divisions at issue (IMD, IBD, and Securities), as Dr. Shaw has done, there are no systematic gender disparities. As shown in Tables 1 and 2 above (*supra* 15-16), when using the same groupings as Dr. Farber for time period and title (*i.e.*, associate and vice president), there are *no disparities for cross-ruffing* at all, *no disparities for 11 of 12 manager quartiling groupings*, and *no disparities for 14 of 18 groupings for 360 reviews*.

There are additional reasons why reasonable jurors would not be swayed by Dr. Farber’s mere calculation of bottom-line, average gender gaps associated with 360 reviews, manager quartiling, and promotions processes aggregated across all three Goldman Sachs Divisions. As Dr. Farber readily concedes, his models are incapable of determining whether there were gender

²¹ Although Judge Francis held that Dr. Farber’s pooled regression models were admissible at class certification based on the limited record at the time (Mar. 10, 2015 Order (ECF No. 363) at 22), Plaintiffs do not and cannot dispute that a “rational juror,” which is the test for summary judgment, *Green v. Town of E. Haven*, 952 F.3d 394, 407 (2d Cir. 2020), could reject Dr. Farber’s pooled models for flunking a Chow test. Indeed, courts routinely reject regression analyses that pool data when a Chow test indicates that pooling is not appropriate. *See, e.g., Florida v. United States*, 885 F. Supp. 2d 299, 325-26 (D.D.C. 2012) (three-judge panel; Garland, Kollar-Kotelly, Huvelle, JJ.) (rejecting as “not methodologically appropriate” an expert’s calculations that inappropriately pooled data from different types of elections “without offering a reason to believe that” pooling was appropriate); *Coates v. Johnson & Johnson*, 756 F.2d 524, 540-42 (7th Cir. 1985) (affirming judgment that pooling year-by-year samples was inappropriate where Chow test so indicated).

disparities in any particular Division, and so his models improperly reflect a classwide, average gender disparity even if an entire Goldman Sachs Division had no gender disparities:

Q. If there were statistically significant disparities with respect to the application of the challenged processes for men and women in the Investment Banking and Investment Management division, but not the Securities division, would your single model be able to identify those disparities?

A. No.

Q. Why not?

A. Well, my model is estimating an average effect *across the company*. . . . And *it might be that in a particular group there wouldn't be an obvious disparity* . . .

(GS Ex. 16 at 85:16-86:11 (Farber Tr.) (emphasis added).)

As Judge Easterbrook explained, an aggregated statistical analysis such as Plaintiffs' of discretionary, subjective employment processes applied across 16 years to dozens of highly skilled and varied jobs, "begs the question," because it merely *assumes* (rather than proves, let alone compels a finding) that the challenged processes actually had a disparate impact across the entire class. *Bolden v. Walsh Const. Co.*, 688 F.3d 893, 896 (7th Cir. 2012). Judge Easterbrook used the following example to illustrate the point: "If [the company] had 25 superintendents, 5 of whom discriminated in awarding overtime, aggregate data would show that black workers did worse than white workers—but that result would not imply that all 25 superintendents behaved similarly." *Id.* Again, Dr. Farber admits that his approach suffers from this exact same fatal flaw, because his approach does not even purport to show that the challenged processes disparately impacted female class members in any particular Division, let alone systemically across the class. (GS Ex. 16 at 85:16-86:11 (Farber Tr.)). This is precisely why the Supreme Court held in *Wal-Mart* that "[m]erely showing that Wal-Mart's policy of discretion has produced an overall sex-based disparity" at the "national level" cannot "raise the inference that a company-wide policy of

discrimination is implemented by discretionary decisions at the store and district level.” *Wal-Mart Stores, Inc. v. Dukes*, 564 U.S. 338, 357 (2011).

This is also why the court in *Abram v. United Parcel Serv. of Am., Inc.*, 200 F.R.D. 424, 431 (E.D. Wis. 2001) rejected the plaintiffs’ statistical analysis that showed, just like Dr. Farber does here, that there was “a statistically significant gap in compensation . . . when the data are considered *in the aggregate*” (emphasis in original). As here, “focusing on the aggregate numbers . . . masks differences from district to district,” which “belies the notion that class members have been affected in common ways.” *Id.*; see also *Kassman v. KPMG LLP*, 416 F. Supp. 3d 252, 283 (S.D.N.Y. 2018) (Schofield, J.) (expert’s “decision to aggregate data across function, thereby equating job and job title, obfuscated the principal explanatory variable”).

Unsurprisingly, Plaintiffs’ cases in which “courts have granted summary judgment” to the plaintiff on the basis of aggregate statistics (Pls.’ Mot. 18) all involved markedly different facts, and serve only to highlight the fatal flaws in Dr. Farber’s oversimplified analysis here. None of Plaintiffs’ cases involved “highly skilled” “professional” employees engaged in a “far-ranging diversity” of jobs requiring different “skills,” making Plaintiffs’ “burden to show that [the company’s] promotion [and assessment] decisions were based on gender—rather than on factors other than . . . sex—especially difficult to meet.” *Valentino v. U.S. Postal Serv.*, 674 F.2d 56, 66-67, 72 (D.C. Cir. 1982) (Ginsburg, J.) (affirming judgment for defendant after bench trial); see *Coser v. Moore*, 739 F.2d 746, 750 (2d Cir. 1984) (“generalized statistical data may be less persuasive . . . when highly specialized jobs are in issue, as opposed to identical positions requiring only general skills”). None of Plaintiffs’ cases involved a challenge to processes even remotely akin to the processes challenged here: subjective and discretionary evaluation and promotion criteria, applied in different ways by individual decision-makers in discrete large businesses across

an extended class period, which requires Plaintiffs to prove something that is “quite unbelievable”—namely, that “all managers [] exercise their discretion in a common way” to the detriment of women in three large Divisions at Goldman Sachs with more than 100 Business Units over a period of 16 years. *Wal-Mart*, 564 U.S. at 356. And, unlike here, in none of those cases did the plaintiffs’ own statistical evidence show only “sporadic” pockets of gender disparities, as that evidence does here, which is fatal to a class claim premised on “systemwide discrimination.” *Valentino*, 674 F.2d at 67. Instead, every case Plaintiffs cite involved *standardized testing criteria* with no elements of discretion (such as a timed running test) and for which it was *actually* undisputed—unlike here—that the testing caused a consistent disparate impact across the class:

- *Easterling v. Conn.*, 783 F. Supp. 2d 323 (D. Conn. 2011): Application of a standardized test involving no discretion (a timed 1.5 mile run) to a single job category (female corrections officers) that caused an undisputed statistically significant disparate impact on women all three times the 1.5 mile run was used in testing.
- *United States v. City of N.Y.*, 637 F. Supp. 2d 77 (E.D.N.Y. 2009) (Garaufis, J.): Application of standardized tests involving no discretion (multiple-choice tests) to a single job category (applicants to FDNY fire academy) that caused an undisputed statistically significant disparate impact on black and Hispanic applicants every time the test was used.
- *NAACP v. N. Hudson Reg’l Fire & Rescue*, 742 F. Supp. 2d 501 (D.N.J. 2010): Application of a standardized hiring criteria involving no discretion (a requirement to live in a particular county) to a single job category (entry-level firefighter positions) that caused an undisputed and consistent statistically significant disparate impact on black applicants.
- *EEOC v. Joint Apprenticeship Comm.*, 164 F.3d 89 (2d Cir. 1998): Application of a standardized exclusion criteria involving no discretion (age cut-off and education level) to three job categories (applicants to three apprenticeship programs) that caused an

undisputed and consistent statistically significant disparate impact on women and black applicants with no pockets or gaps in the data.

- *United States v. City of Warren*, 759 F. Supp. 355 (E.D. Mich. 1991): Application of a standardized hiring criteria involving no discretion (requirement to live in a particular city) to two job categories (firefighter and police officer positions), which resulted in not a single black person ever being hired by the city (*i.e.*, the inexorable zero).
- *Kirkland v. N.Y. State Dep't of Corr. Servs.*, 520 F.2d 420 (2d Cir. 1975): Application of a standardized test for promotions involving no discretion (a written, multiple-choice test) to a single job category (corrections officers) that caused an undisputed and “uniform[]” statistically significant disparate impact on minorities “at all of the State’s facilities.” *Id.* at 425.

Plaintiffs also are wrong that these decisions give this Court license to weigh the evidence and “reject[]” Goldman Sachs’ “statistical criticisms” of Dr. Farber’s model. (Pls.’ Mot. 20.) Weighing evidence is expressly prohibited in resolving Plaintiffs’ motion for summary judgment. *See Gibson*, 892 F.2d. at 1132 (“in a Title VII case particularly,” “[i]t is not the trial court’s function to weigh the evidence and resolve the factual issues”). The courts did “not [] weigh the evidence” in any of Plaintiffs’ cases ruling on summary judgment. *NAACP*, 742 F. Supp. 2d at 510. Rather, because the evidence of disparate impact was so overwhelming, the defendants could not, and did not, seriously dispute it in opposing summary judgment, but instead conceded disparate impact or “resort[ed] to abstract arguments” about “Plato’s ‘allegory of the cave’” to criticize “statistical testing in general.” *City of N.Y.*, 637 F. Supp. 2d at 76, 98 (“[T]he City has conceded the accuracy of the calculations of Plaintiffs’ experts, which provide ample support for the statistical and practical significance of the disparities at issue.”).²²

²² *See also Easterling*, 783 F. Supp. 2d at 333 (“defendant’s sole expert on the issue of statistical disparity . . . concluded that [plaintiff’s expert’s] calculations were accurate”); *NAACP*, 742 F. Supp. 2d at 519 (“both [parties’] experts’ reports support a statistical finding of disparate impact”); *City of Warren*, 759 F. Supp. at 357-58 (“Warren has expressly stated that, for purposes of its

b. Even accepting Dr. Farber’s models at face value, Dr. Farber fails to show that the challenged processes caused any purported gender disparities.

Even if Dr. Farber’s regression models were accepted at face value as undisputed, they would fail to make out a *prima facie* case of disparate impact because they do not establish that the challenged processes caused any purported gender disparities. Plaintiffs’ case rests on the theory that the 360 review, manager quartiling, and promotion processes were “uniform” throughout the 16-year class period (SAC ¶¶ 37, 40, 68), yet Dr. Farber’s own models found that there was *no* discrimination against female professionals in any Division during 2016, 2017, and 2018 in either 360 reviews or manager quartiling. (GS 56.1 Resp. ¶¶ 25, 33, Pls.’ Ex. 18 ¶ 114 Tables 19, 20 (Farber Rpt.).) Likewise, Dr. Farber opined that class members were not subject to discrimination with respect to promotions in 2002, 2003, 2010, 2012, and 2015 (30% of the class period)—in fact, more women were promoted to managing director in those years than Dr. Farber’s model predicted. (GS 56.1 Resp. ¶ 58, Pls.’ Ex. 18 ¶ 152, Table 28 (Farber Rpt.); Pls.’ Ex. 58 ¶ 117, App. A Table 28 (Farber Reply Rpt.).) As explained in Goldman Sachs’ summary judgment motion, the sporadic and inconsistent “impact” Dr. Farber found refutes any notion that “a uniformly applied, common practice is driving the [allegedly] unequal outcomes at issue” for any groups or time periods. *Richardson*, 2021 WL 1910689, at *9; *see* GS SJ Br. 3-6, 23-27.²³

present motion and its defense against the Government’s motion, it is not arguing ‘that the residency requirements did not have an adverse impact against blacks.’” (quoting defendant’s brief)); *see also Kirkland*, 520 F. 2d at 425 (decided on appeal *after trial*, and defendants were “not disputing the accuracy of the[] figures” concerning disparate impact presented by plaintiffs).

²³ Plaintiffs try to bolster Dr. Farber’s flawed statistics by pointing to an “internal” Goldman Sachs document that supposedly identifies “gender gaps.” (Pls.’ Mot. 10.) But Plaintiffs acknowledge in a footnote that this document has no probative value because it “does not address whether the gender gaps are statistically significant.” (*Id.* at n.43.) *See Walker v. City of N.Y.*, 2014 WL 1244778, at *8 n.8 (S.D.N.Y. Mar. 26, 2014) (Failla, J.) (“raw numbers” are “meaningless” and

Plaintiffs attempt to address Dr. Farber’s patchwork of findings by moving for summary judgment as to 360 reviews and quartiling only “through the end of 2015” (Pls.’ Mot. 1), despite alleging that these processes were used after that time period with no disparate effect on women. (Pls.’ 56.1 ¶ 1.) But that contrivance does not make Dr. Farber’s own findings go away. The Second Circuit has rejected this exact kind of ‘don’t look over here’ argument, holding that a plaintiff’s attempt to “isolat[e] a few work-groups and analyz[e] the effect” of a company-wide reduction-in-force process based on just those cherry-picked “work-group[s] is misleading at best.” *Smith*, 196 F.3d at 370. “[A] plaintiff may not ‘gerrymander’ data to skew the results of statistical analyses in her favor”; a “decision-making process either caused [a] disparate impact or it did not,” and “[h]aving chosen” to stake their claim on a supposedly “identical” “overall process,” “[i]t is only reasonable to infer a disparate impact from the . . . process if all persons who were subject to the process” experienced such impact. *Id.* at 369-70.²⁴

cannot “support a prima facie case” of discrimination). And Dr. Farber himself has acknowledged that “no conclusions” can be drawn from analyses such as the one cited by Plaintiffs which do not account for differences between professionals relevant to their performance, such as tenure, education, or production metrics. (GS 56.1 Resp. ¶ 41, Pls.’ Ex. 56 at -257, -272 (GS0918243); Pls.’ Ex. 57 at -438 (GS0176436); Pls.’ Ex. 18 ¶¶ 71, 77 (Farber Rpt.).)

²⁴ Plaintiffs also bizarrely claim that Dr. Farber’s aggregate statistics prove that Goldman Sachs’ 360 review process discriminated against women in 2002, even though Goldman Sachs did not begin to implement firm-wide 360 reviews *until 2003*, and did not fully implement them in two of three at-issue Divisions until 2004. (GS 56.1 Resp. ¶ 5, GS Ex. 7 at 38:14-19 (9/5/13 Landman Tr.).) Plaintiffs cannot cure Dr. Farber’s absence of data by mischaracterizing Judge Francis’ September 10, 2012 Order. (Pls.’ Mot. 9 n.33.) That Order merely provided that Goldman Sachs did not need to produce information from an older version of its PeopleSoft database—which *has never contained 360 review data*, because 360 review data is not stored in the PeopleSoft database but rather in the Firmwide Review System (*see* GS Ex. 27 ¶ 2 (Obradovich Aff.); GS Ex. 28 ¶¶ 1-2 (Pathak Aff.))—provided that Goldman Sachs “agree[d] not to argue *for purposes of class certification* that information in the older *PeopleSoft database* is materially different from that in the current database.” (Sept. 10, 2012 Order (ECF No. 159) at 40 (emphasis added).) Nothing in that Order says anything about 360 review data, much less suggests that Goldman Sachs somehow conceded liability for a non-existent process or based on data that does not exist.

Without any facts or law to support their motion, Plaintiffs' motion for summary judgment should be denied, and the Court should enter summary judgment for Goldman Sachs.

2. Dr. Shaw does not “agree” that the challenged processes caused gender disparities.

Because their statistics fall far short of demonstrating that the challenged processes caused a disparate impact, Plaintiffs misrepresent Dr. Shaw's opinions in hopes of convincing the Court that “both experts' models [find] . . . a statistically significant pay gap caused by the challenged practices.” (Pls.' Mot. 21.) While this is simply not true, it does reveal the desperate lengths Plaintiffs will go to try to salvage their claims.

First, Plaintiffs falsely claim that Dr. Shaw somehow “concede[d] that there is a statistically significant gender pay gap caused by the Challenged Employment Practices.” (Pls.' Mot. 3.) Dr. Shaw did no such thing. Her opinions are the exact opposite:

- “The statistical evidence *rejects the claim* that there is a common, systemic gender gap with respect to the 360 review process, the assignment of manager quartiles, or promotion to EMD at GS during the relevant period.” (GS 56.1 Resp. ¶ 51, Pls.' Ex. 17 ¶ 29 (Shaw Rpt.) (emphasis added); *see also id.* ¶¶ 185-238.)
- “The statistical evidence *rejects the claim* that alleged gender gaps in the challenged processes cause an alleged gender gap in compensation.” (*Id.* (emphasis added); *see also id.* ¶¶ 240, 277-96.)

Despite Dr. Shaw's unequivocal opinions, Plaintiffs persist in claiming that Exhibit 68 to Dr. Shaw's report reflects her “agree[ment] that the 360 Review and Manager Quartiling practices result in lower compensation for women.” (Pls.' Mot. 12.) That is a blatant misrepresentation of Exhibit 68, which in any case has nothing to do with the certified claims. As explained above (*supra* 16-18), Exhibit 68 reflects Dr. Shaw's analysis of differences in compensation between male and female professionals *using Dr. Farber's models* with additional corrections, solely for

the purpose of refuting Dr. Farber’s claim of gender disparities in compensation *that are unrelated to the challenged processes*. (GS 56.1 Resp. ¶ 51, Pls.’ Ex. 17 ¶ 241 (Shaw Rpt.).) Accordingly, nothing in Exhibit 68 has anything to do with whether the 360 review, quartiling, and cross-ruffing processes caused gender disparities. In fact, as shown in Table 3 above, Dr. Shaw’s median regression shows that there are no significant gender gaps in compensation for any group of class-relevant employees (*supra* 17).

Second, Plaintiffs claim that Dr. Shaw’s opinions must be rejected because Dr. Shaw supposedly claims that Plaintiffs “must show that the challenged processes have the same effect on every Class member.” (Pls.’ Mot. 21.) Dr. Shaw never said this, nor does she contend that “it is not enough to show a statistical pattern of average gender differences.” (*Id.*) Like Dr. Farber, Dr. Shaw also calculated average gender disparities. The difference is that, for each title and time period, Dr. Farber improperly calculated a single, bottom-line average gender disparity across all three large and diverse Divisions, while Dr. Shaw calculated “average gender differences” for each Division. In other words, instead of calculating a single average like Dr. Farber, Dr. Shaw calculated three averages—and controlled for relevant factors, such as employee production, that Dr. Farber excluded. (Pls.’ Ex. 17 ¶¶ 190, 209-41 (Shaw Rpt.).) Thus, Dr. Shaw’s models do precisely what Plaintiffs say they should do: analyze “average effects on a protected group.” (Pls.’ Mot. 21.)²⁵ There is no basis to disregard her findings when considering Plaintiffs’ summary judgment motion.

²⁵ As this Court explained in certifying the class, “Plaintiffs must first show that Defendants used ‘a common mode of exercising discretion.’” (Class Cert. Order 25.) Unlike Dr. Farber, who admittedly “did not examine” whether any challenged “process was applied in a different way within a particular division versus another division” (GS Ex. 16 at 56:18-24 (Farber Tr.)), Dr. Shaw did—and found that “the evidence is clear that the processes function differently *at least at*” the Division level (Pls.’ Ex. 17 ¶ 27 (Shaw Rpt.)), meaning that the processes should be analyzed at least at that level.

B. Plaintiffs’ evidence as to 360 reviews and promotions fails to make out a *prima facie* case for the independent reason that Plaintiffs have no justification for their decision to lump together the discrete components of those processes.

Plaintiffs contend that Dr. Farber’s statistical analysis of the *overall* 360 review and promotions processes entitles them to summary judgment on the *prima facie* element of their disparate impact claim. (Pls.’ Mot. 16-24.) In making this argument, Plaintiffs assert that under Title VII, they may analyze the overall processes, as opposed to the distinct components of each, but that is not the case. For this reason alone—because Plaintiffs cannot merge the distinct elements of either the overall 360 review or the promotions processes—they are not entitled to summary judgment as to those processes, which instead should be granted to Goldman Sachs. And certainly Plaintiffs cannot manufacture “evidence” supporting their cross-ruffing claim through Dr. Shaw’s response to Dr. Farber’s inappropriate analysis of overall promotions, which shows no disparate impact either in cross-ruffing specifically or promotions generally.

1. Plaintiffs’ failure to analyze the separate steps of the 360 review and promotion processes precludes summary judgment in their favor and requires summary judgment for Goldman Sachs.

As an initial matter, regarding promotions, Plaintiffs sought and obtained certification of only the “cross-ruffing” element. (Class Cert. Order at 28.) Plaintiffs have introduced no competent evidence to establish discriminatory impact in cross-ruffing, foreclosing summary judgment for them on their Title VII claim (and their NYCHRL claim, *see infra* 40-41). By contrast, the only evidence in the record (Dr. Shaw’s) is that cross-ruffing is not discriminatory, and does not lead to discriminatory promotions outcomes, entitling Goldman Sachs to summary judgment on Plaintiffs’ cross-ruffing claims.

But there is another reason why Plaintiffs were legally required to analyze the component parts of all three class-certified processes (360 reviews, quartiling and cross-ruffing). In 1991, Congress amended Title VII to expressly require plaintiffs asserting disparate impact claims to

identify the “particular employment practice” that they contend caused a disparate impact. As revised, Title VII permits plaintiffs to group together discrete “elements” of a practice and analyze them as one only where plaintiffs “can demonstrate to the court that the elements of a respondent’s decisionmaking process are not capable of separation for analysis.” Civil Rights Act of 1991, Pub. L. No. 102-166, § 105(a), 105 Stat 1071 (codified at 42 U.S.C. § 2000e-2(k)(1)(B)).

As explained at length in Goldman Sachs’ summary judgment brief, this is a critical feature of Title VII. (*See* GS SJ Br. 20-23.) Assessing the effect of an employer’s overall “decisionmaking process” is a narrow exception to the rule that Plaintiffs are “responsible for isolating and identifying the *specific* employment practices that are allegedly responsible for any observed statistical disparities.” *Smith v. City of Jackson*, 544 U.S. 228, 241 (2005). The Supreme Court has been crystal clear that “it is not enough to simply allege that there is a disparate impact on workers, or point to a generalized policy that leads to such an impact.” *Id.* Rather, Plaintiffs “must make some effort to isolate any allegedly adverse practices and to examine their individual effects” or show that an overall process’s “many steps were so intertwined that they were not capable of separation for analysis.” *Davis v. Cintas Corp.*, 717 F.3d 476, 497 (6th Cir. 2013) (alterations omitted). Because Plaintiffs did not even attempt to carry their burden of separately analyzing the indisputably discrete elements of the 360 review and promotions processes or showing that they are “not capable of separation for analysis,” 42 U.S.C. § 2000e-2(k)(1)(B), summary judgment cannot be granted in their favor.

Plaintiffs’ contrary argument that “there is no reason to divide a challenged process into subparts for analysis” (Pls.’ Mot. 23) ignores the plain language of the 1991 amendments to Title VII. The only authorities Plaintiffs cite in support are from the 1970s—nearly 20 years before Congress changed the law to expressly require a plaintiff to “divide a challenged process into

subparts for separate analysis.” (*Id.*) It is completely irrelevant that EEOC guidance from 1979 found it permissible to examine “the combined effect of all selection procedures leading to the final employment decision,” EEOC-NYVTA-1979-1, *Questions & Answers to Employee Selection Procedures* (1979), or that in 1975 the Second Circuit believed there to be “little relevance” in analyzing separately “the five sub-tests comprising” a challenged employment test, *Kirkland*, 520 F.2d at 425. Congress disagreed with those views, and enacted a statute making clear that Plaintiffs may only analyze discrete elements of a process as one if they “demonstrate . . . the elements of a respondent’s decisionmaking process are not capable of separation for analysis.” 42 U.S.C. § 2000e-2(k)(1)(B). The Court in *EEOC v. Freeman*, 961 F. Supp. 2d 783, 900-01 (D. Md. 2013), rejected the same overture to error that Plaintiffs invite here, granting summary judgment *for defendant* despite the EEOC’s argument—based on decisions that “all predate the Congressional codification of the disparate impact particularity requirement in 1991”—that it was not “necessary” to analyze separately each component of “a multi-faceted, multi-step policy.”²⁶

2. Dr. Shaw’s analysis of cross-ruffing according to statutory standards and the Certification Order cannot salvage plaintiffs’ summary judgment motion.

Even if the Court were to conclude that Dr. Farber properly analyzed overall promotions rather than the certified cross-ruffing process (which would be legally erroneous), that would not be grounds for “grant[ing] summary judgment on Plaintiffs’ *prima facie* case of disparate impact

²⁶ The fact that the Court certified a class with respect to the 360 review, quartiling, and cross-ruffing processes does not relieve Plaintiffs of their burden at the merits stage to comply with Title VII’s express requirement to pinpoint the particular “element” of the defendant’s “decisionmaking process” that caused a disparate impact. 42 U.S.C. § 2000e-2(k)(1)(B)(i). *See Grant v. Metro. Gov’t of Nashville & Davidson Cty.*, 446 F. App’x 737, 740 (6th Cir. 2011) (notwithstanding district court’s certification of a class, reversing judgment after bench trial where plaintiffs “never attempted to demonstrate that the elements of th[e] [promotions] process are incapable of separation for analysis”).

in promotions.” (Pls.’ Mot. 24.) In support of this argument, Plaintiffs try to flip the burden onto Goldman Sachs, contending that they are entitled to summary judgment because Goldman Sachs’ expert, Dr. Shaw, supposedly did not “justif[y]” her decision to analyze separately the “nomination” and “cross-ruffing” components of the promotions process. (*Id.*) According to Plaintiffs, this means that the Court should “reject Dr. Shaw’s” analysis of cross-ruffing, thereby supposedly leaving Dr. Farber’s analysis of promotions unchallenged. (*Id.*)

This argument is frivolous. Dr. Farber’s faulty analysis of overall promotions is not unchallenged. In response to Dr. Farber’s examination of overall promotions, Dr. Shaw *did analyze* overall promotions as well, finding that “women fare as well as or better than men.” (GS 56.1 Resp. ¶ 58, Pls.’ Ex. 17 ¶¶ 35, 239 (Shaw Rpt.).) And even taking Dr. Farber’s flawed methodology and his improper analysis of promotions generally standing alone, it does not make out a *prima facie* case of disparate impact. Dr. Farber found that class members were not subject to discrimination with respect to promotions in 2002, 2003, 2010, 2012, and 2015 (30% of the class period). (GS 56.1 Resp. ¶ 61, Pls.’ Ex. 18 ¶ 152, Table 28 (Farber Rpt.), Pls. Ex. 58 ¶ 117, App. A Table 28 (Farber Rebutl. Rpt.).) As explained in Goldman Sachs’ motion for summary judgment (ECF No. 1243 at 3-5, 23-27), and as discussed above (*supra* 29-31), the absence of common impact across the class members and across the class period refutes the notion that the processes *themselves* caused any gender disparities identified by Dr. Farber’s analysis (as opposed to other factors such as random individual manager discretion, or unmeasured factors).

In any event, Dr. Shaw’s decision to analyze cross-ruffing was fully justified.

First, as already explained, Plaintiffs, not Defendants, bear the burden of demonstrating that the discrete components of the promotions process “are not capable of separation for analysis.” 42 U.S.C. § 2000e-2(k)(1)(B)(i); *see id.* (“the complaining party shall demonstrate . . .”); *Campbell*

v. *Nat'l R.R. Passenger Corp.*, 311 F. Supp. 3d 281, 326 (D.D.C. 2018) (“[P]laintiffs have failed to meet their burden to show that Amtrak’s selection procedures are ‘not capable of separation for analysis.’”). That failure by Plaintiffs to comply with the statutory requirements means that Goldman Sachs is entitled to summary judgment, not the other way around.

Second, the fact that the Court certified only cross-ruffing and not promotions generally (Class Cert. Order at 28) provided Dr. Shaw with ample “justification” (Pls.’ Mot. 24) for analyzing cross-ruffing specifically.²⁷ It is Plaintiffs who must justify Dr. Farber’s decision to analyze the promotions process overall, when its component parts clearly are “capable of separation for analysis” (as Dr. Shaw showed), and when Plaintiffs themselves conceded in an interrogatory response that their claims were based on cross-ruffing specifically, not “promotions” writ large. (*See supra* 7.)²⁸

In sum, because Plaintiffs made no effort to carry their burden of justifying Dr. Farber’s decision to lump together the component parts of 360 reviews and promotions, Plaintiffs’ Title VII disparate impact claim as to those processes is dead in the water, and Goldman Sachs is entitled to summary judgment on its cross-motion. *See, e.g., Campbell*, 311 F. Supp. 3d at 326-27 (granting summary judgment to employer where plaintiffs “failed to meet their burden to show that

²⁷ Plaintiffs also contend that analyzing cross-ruffing “artificially reduces the statistical power of the analysis.” (Pls.’ Mot. 24.) But there is nothing artificial about limiting the analysis to the “cross-ruffing” process that the Court certified. Plaintiffs also offer nothing to substantiate their conjecture that the pool of professionals nominated for cross-ruffing was too “small[] . . . to study.” (Pls.’ Mot. 24.) That criticism is irrelevant in any event because the Second Circuit has expressly rejected efforts to dilute Title VII’s legal requirements for statistical proof just because a properly conducted analysis poses challenges to the plaintiff. *See Coser*, 739 F.2d at 750 (statistical analysis must reflect “decentralized” hiring and smaller “availability pool” of qualified candidates, even if this leaves the court “less confident of the validity of statistical proof”).

²⁸ As Dr. Farber testified: “I did not do an analysis of cross-r[uff]ing. . . . Cross-r[uff]ing was part of the promotion process, but it’s not the entire promotion process. And I analyzed the overall effect of the promotion process, not cross-r[uff]ing specifically.” (GS Ex. 16 at 134:15-16, 149:12-16 (Farber Tr.).)

[employer's] selection procedures [we]re 'not capable of separation for analysis'"); *Davis*, 717 F.3d at 497-98 (affirming summary judgment to employer because plaintiff "did not explain why the well-defined, discrete elements of the Meticulous Hiring System are 'not capable of separation for analysis'"); *Grant*, 446 F. App'x at 740 (to the same effect).

II. Plaintiffs have not established a *prima facie* case of disparate impact under the NYCHRL.

Although Plaintiffs are correct that the NYCHRL differs from Title VII in certain respects (Pls.' Mot. 25), Plaintiffs acknowledge that the NYCHRL, just like Title VII, requires them to "prov[e] that the Challenged Employment Practices *cause* an adverse impact on the Class." (*Id.* at 3 (emphasis added)); *see* N.Y.C. Code § 8-107(17)(a)(1) (plaintiff must demonstrate that "a policy or practice of [defendant] or a group of policies or practices of [defendant] *results in* a disparate impact to the detriment of [a protected group]") (emphasis added). Accordingly, Plaintiffs' failure of proof on causation is fatal to both their Title VII claim and their NYCHRL claim. Plaintiffs' arguments to the contrary are meritless.

First, Plaintiffs contend that courts "constru[e] NYCHRL's provisions broadly in favor of discrimination plaintiffs." (Pls.' Mot. 25.) But that broad construction principle does not relieve Plaintiffs from proving indispensable elements of their NYCHRL claim (such as causation), and it "does not alter the kind, quality or nature of evidence that is necessary to support or defeat a motion for summary judgment under Rule 56." *Arnou v. Aeroflot Russian Airlines*, 980 F. Supp. 2d 477, 486 (S.D.N.Y. 2013) (Forrest, J.); *see Fitchett v. City of N.Y.*, 2021 WL 964972, at *24 (S.D.N.Y. Mar. 15, 2021) (Engelmayer, J.) (granting summary judgment to defendants "[e]ven under [the NYCHRL's] more permissive standards" because plaintiff "has not identified a facially neutral policy that, with the data subjected even to minimal rigor, has yielded racially disparate promotion outcomes"); *Assistant Deputy Wardens/Deputy Wardens Ass'n v. City of N.Y.*, 2019

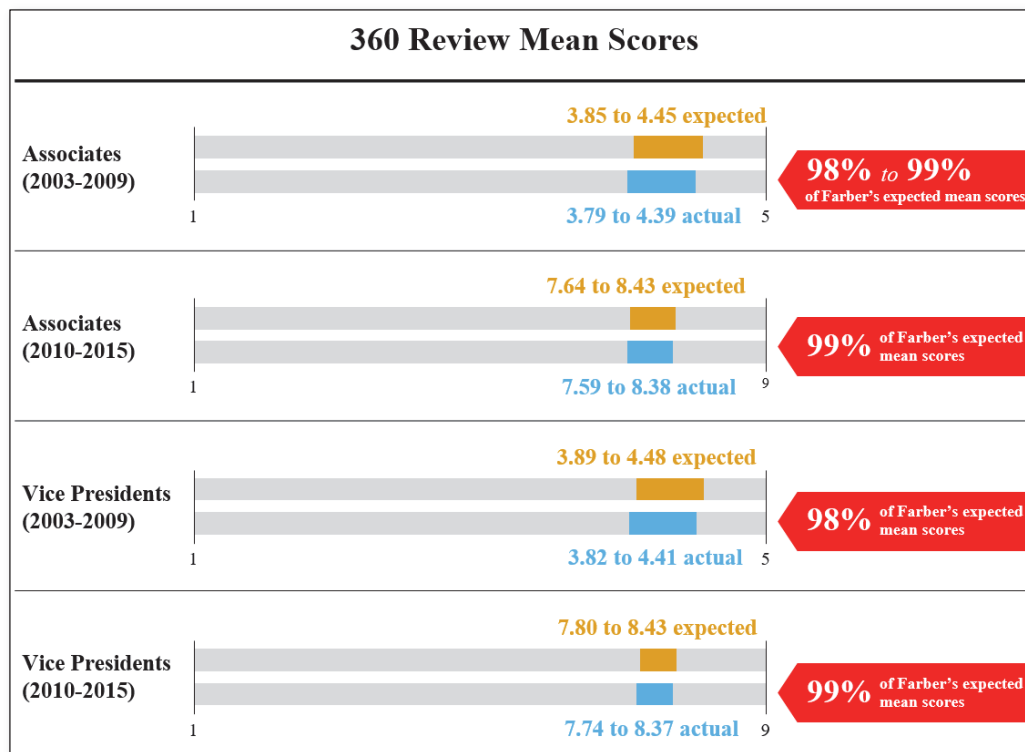
WL 4015119, at *2 (E.D.N.Y. Aug. 26, 2019) (Block, J.) (granting summary judgment to defendants on Title VII and NYCHRL claims, noting that “[t]he more [] lenient standard does not change that analysis”).

Tellingly, Plaintiffs’ sole authority for the proposition that the NYCHRL’s “more lenient” standard requires summary judgment for Plaintiffs (Pls.’ Mot. 27) *granted summary judgment to defendants*. In that case, the plaintiff claimed that New York City’s policy of terminating provisional emergency medical technicians who failed a driving skills test had a disparate impact on women over age 40. *Teasdale*, 2013 WL 5300699, at *9. The court held that plaintiff’s statistical analysis “ha[d] no probative value,” leaving only defendants’ evidence, which showed that “the pass rate ratio for women is [] 91% of the rate for men and the pass rate ratio for candidates over 40 is also 91% of the rate for candidates under 40 years of age.” *Id.* at *9-10. Based on this evidence, the court granted summary judgment to defendants, holding “that plaintiff has failed to show through statistical analysis or any other evidence that the [driving test] results in a disparate impact to any protected group.” *Id.* at *13.

Teasdale is on all fours with this case, although not for the reason Plaintiffs suggest. Even under Dr. Farber’s flawed model, the gender differences between women’s actual and expected 360 review scores, top manager quartile placements, and promotions are in the same range as or even less than those deemed insufficient to create a triable issue in *Teasdale* under *both* Title VII and the NYCHRL. Dr. Farber’s faulty promotions model predicted that 3.68% of eligible women should have been promoted from vice president to managing director, and that the actual rate of promotion was 3.27%, **which is 89% of Dr. Farber’s expected promotion rate**. (GS 56.1 Resp. ¶ 58, 64, Pls.’ Ex. 58 App. A, Table 24 (Farber Rebt. Rpt.).) Dr. Farber’s model predicted that 29% of female associates and vice presidents should have been placed in the first quartile and that

in fact 26% were so placed, **which is 88% of his expected rate.** (GS 56.1 Resp. ¶ 33, Pls.’ Ex. 58 App. A, Tables 5-8 (Farber Rebtl. Rpt.).) And, as shown in the table below, Dr. Farber found even smaller disparities for the 360 review process:

Table 4



(GS 56.1 Resp. ¶ 25, Pls.’ Ex. 58 App. A, Tables 9-12 (Farber Rebtl. Rpt.).)

Second, Plaintiffs get no help from the fact that the NYCHRL permits Plaintiffs to challenge a “group of policies or practices [that] results in a disparate impact.” (Pls.’ Mot. 26-27 (emphasis in original).) According to Plaintiffs, “[u]nder New York law, once impact has been established, the plaintiff need not show which policies establish impact.” (*Id.*) But this argument fails because Dr. Farber’s findings of sporadic and inconsistent impact refute the notion that *any* of three challenged processes, either alone or in some combination, caused a disparate impact.

Finally, regarding cross-ruffing, that process—and not promotions generally—was certified for class treatment. (Class Cert. Order at 28.) Whatever the scope of the NYCHRL, it

does not relieve Plaintiffs of the class definition or the class certification Order. Because Plaintiffs failed to submit any evidence in support of their cross-ruffing theory, they are not entitled to summary judgment on that issue, which should be granted to Goldman Sachs.

CONCLUSION

The Court should deny Plaintiffs' motion for summary judgment and instead enter summary judgment for Goldman Sachs.

Respectfully,

Barbara B. Brown (*admitted pro hac vice*)
Carson H. Sullivan (*admitted pro hac vice*)
PAUL HASTINGS LLP
875 15th Street, N.W.
Washington, D.C. 20005

Patrick W. Shea
PAUL HASTINGS LLP
200 Park Avenue
New York, New York 10166

/s/ Robert J. Giuffra, Jr.
Robert J. Giuffra, Jr.
Sharon L. Nelles
Ann-Elizabeth Ostrager
Hilary M. Williams
SULLIVAN & CROMWELL LLP
125 Broad Street
New York, New York 10004
(212) 558-4000

Amanda Flug Davidoff
SULLIVAN & CROMWELL LLP
1700 New York Avenue, N.W., Suite 700
Washington, D.C. 20006

*Attorneys for Defendants
Goldman Sachs & Co. and
The Goldman Sachs Group, Inc.*

October 12, 2021